基于横纵向联合控制的多目标优化车辆跟驰研究 *

李孟凡，秦文虎†，云中华

(东南大学 仪器科学与工程学院，南京 210096)

摘　要： 为解决车辆在拥堵环境中因车速波动较大所带来的跟驰平稳性较差，跟踪无效或不安全等问题，提出了基于车辆模型和深度强化学习的多目标优化跟驰方案。首先基于车辆横纵向动力学建立车辆跟驰模型，然后根据车间距误差、速度误差、横向偏差及相对偏航角等，利用深度确定性策略梯度算法得到跟驰车的加速度和转向角，以更平稳安全地控制跟驰车辆。经 NGSIM 公开驾驶数据集进行测试与验证，该方案可有效地提升跟驰车辆的稳定、舒适与安全性，对保证交通安全和提升道路通行能力具有重要意义。

关键词： 车辆跟驰；横纵向联合控制；深度确定性策略梯度；NGSIM

中图分类号： TP391　　doi: 10.19734/j.issn.1001-3695.2021.12.0687

Multi-objective optimal car-following model with lateral and longitudinal control

Li Mengfan, Qin Wenhu†, Yun Zhonghua

(*Southeast University*, *School of Instrument Science&Engineering*, *Nanjing* 210096, *China*)

Abstract: Due to the large fluctuations and poor stability of the vehicle in the congested environment, it is difficult to track the vehicle in front of it in time. This paper developed a multi-objective optimal car following algorithm. The algorithm used a vehicle lateral and longitudinal dynamic model and deep reinforcement learning to establish a car-following model. Then, according to the vehicle distance difference, speed difference, lateral deviation, and relative yaw angle, the deep deterministic strategy gradient algorithm obtained the acceleration and steering angle of the following vehicle. Therefore, the algorithm can control the following car more smoothly and safely. After testing and verification with the NGSIM driving data set, the results show that this method can effectively improve the stability, comfort, and safety of car following. It can promote traffic safety and improve road capacity.

Key words: car following; lateral and longitudinal; deep deterministic policy gradient; NGSIM

0　引言

车辆跟驰是重要的自动驾驶辅助技术，能够减轻驾驶员的驾驶负担，提高驾驶舒适性，减少交通安全事故。但在交通拥堵状况中由于车辆频繁的加减速，跟驰效果较差，因此低速跟驰研究得到了广泛关注[1]。

以往的自动驾驶跟驰模型多为理论驱动，主要在车辆跟驰行为基础上，以数学、物理学模型表达跟驰过程中的各种状态，建立符合驾驶经验的车辆跟驰模型。首个跟驰模型由pipes[2]提出，其假设跟驰车速度与车间距成正比，根据车间距决策跟驰车速度，建立车辆跟驰模型。之后，基于交通流异质性、人类因素、道路因素等陆续提出了基于安全距离、心理生理、刺激反应以及元胞自动机等车辆跟驰模型，但理论驱动跟驰模型难以综合考虑上述各种影响因素，模型预测精度较差，在描述复杂车辆跟驰行为时不够准确。

得益于智能交通的发展，大规模高精度车辆轨迹数据为基于数据驱动的车辆跟驰提供了研究基础，通过对车辆轨迹数据统计分析，挖掘出有关驾驶行为规律，建立对应的拟合关系，从而实现车辆的有效跟驰[3]。目前，数据驱动车辆跟驰模型有基于模糊逻辑、支持向量回归、人工神经网络、深度强化学习模型等。其中深度强化学习跟驰模型[4]是近年的研究热点，如卷积神经网络、递归神经网络、长短时记忆神经网络等方法逐步应用于车辆跟驰研究中。Zhu Meixin 等人[5]利用深度强化学习得到了车辆跟驰策略，建立了从跟驰速度、相对速度、车间距到跟驰加速度的仿人映射模型；Pan Feng 等人[6]基于逆强化学习对真实驾驶数据分析后，获取了人类驾驶员的驾驶特征，设计了奖励函数，得到了更加自然的跟驰驾驶行为；朱冰[7]基于近端策略优化算法建立了车辆跟驰控制策略及基于双前车跟驰结构的主车纵向控制架构，实现了车辆跟驰控制。模型预测控制[8](Model Predictive Control，MPC)多应用于跟驰场景中，重庆大学胡晓松等人[9]开发了一种基于 MPC 的控制器来优化车速和发动机扭矩，在确保跟驰安全的同时实现更好的燃油经济性和更少的废气排放。Mao Jin 等人[10]基于 MPC 算法，设计了可以在线更新权重系数的多目标优化方法，实现了更好的跟驰跟踪性和稳定性。与其他模型相比，深度强化学习车辆跟驰模型能够通过不断地学习适应不同地驾驶环境，具有更好的泛化能力，有助于开发与人类驾驶行为更相似的自主驾驶算法和交通流模型。

驾驶员在跟车行驶时，意识到车辆偏离期望轨迹时，通常会通过降低速度和增大跟车距离，以减小横向控制风险和纵向事故风险[11]。车辆在曲率变化的道路上行驶时，若模型对横向控制不足，会影响车辆的操纵稳定性。通常的跟驰研究大多考虑了纵向加速度决策而忽略了横向路径跟踪，且主要集中于对人类驾驶行为的模拟而忽略了对驾驶行为的优化。本文基于车辆横纵向联合控制并考虑安全性、舒适性，利用深度强化学习算法得到同时决策加速度和转向角的车辆跟驰模型，实现车辆有效地跟随前车行驶。

收稿日期： 2021-12-27；**修回日期：** 2022-02-28　**基金项目：** 江苏现代农业产业关键技术创新资助项目 CX(20)2013，江苏省重点研发计划资助项目(BE2019311)

作者简介： 李孟凡(1995-)，女，江苏连云港人，硕士/硕士研究生，主要研究方向为智能驾驶、汽车安全、仪器仪表；秦文虎(1969-)，男(通信作者)，江苏泰州人，教授，博导，主要研究方向为汽车电子、传感器测控技术、虚拟现实(qinwenhu@seu.edu.cn)；云中华(1983-)，男，河南洛阳人，博士研究生，主要研究方向为仪器仪表、传感器技术、信号处理.

1　车辆动力学模型

在跟驰车辆的运动控制中，首先需建立车辆动力学模型，该研究根据转向和牵引系统，简化转向系统，直接输入前轮转角作为跟驰车方向盘转角，基于车辆横向运动和横摆运动以及纵向运动，建立车辆三自由度动力学模型如下[12]：

$$\begin{cases}\dot{v}_x = v_y\dot{\varphi} + (F_{xf}\cos\delta - F_{yf}\sin\delta + F_{xr})/m \\ \dot{v}_y = -v_x\dot{\varphi} + (F_{xf}\sin\delta + F_{yf}\cos\delta + F_{yr})/m \\ \ddot{\varphi} = ((F_{xf}\sin\delta + F_{yf}\cos\delta)l_f - F_{yr}l_r)/I_z\end{cases} \tag{1}$$

图 1 为车辆动力学模型，图中 XOY 是地面参考坐标系，xoy 是车辆坐标系。v_x, v_y 是车辆质心的纵向和侧向速度，φ 是车辆横摆角，$\dot{\varphi}$ 是车辆横摆角速度，δ 为前轮转向角；l_f 是汽车质心到前轮的距离，l_r 是汽车质心到后轮的距离；l_z 是车辆绕铅垂轴的转动惯量；F_{xf} 和 F_{yf} 分别是前轮所受到纵向力和侧向力，F_{xr} 和 F_{yr} 分别是后轮所受到的作用力。

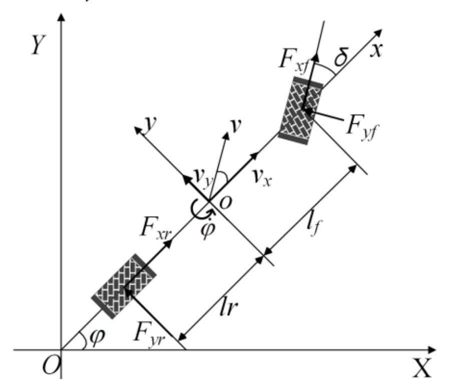


图 1　车辆动力学模型

Fig. 1　Vehicle dynamics model

前后轮胎的侧向力与侧偏角近似成线性关系[13]。式(2)中 α_f，α_r 分别为前轮侧偏角和后轮侧偏角；C_f、C_r 分别为车辆前轮胎转弯刚度和后轮胎转弯刚度[12]。

$$\begin{cases}F_{yf} = 2C_f\alpha_f = 2C_f(\delta - \dfrac{l_f\dot{\varphi} + v_y}{v_x}) \\ F_{yr} = 2C_r\alpha_r = 2C_r\dfrac{l_r\dot{\varphi} - v_y}{v_x}\end{cases} \tag{2}$$

由式(1)(2)即可得到车辆的纵向加速度 $\dot{v}_x$、横向加速度 $\dot{v}_y$ 和横摆角加速度 $\ddot{\varphi}$。

2　基于 DDPG 车辆跟驰决策算法

深度确定性策略梯度算法具有深度神经网络的特征提取能力和强化学习的决策优势，且适用于具有离散输入、连续输出的跟驰决策问题，因此，本文基于深度确定性策略梯度算法建立了车辆跟驰总体策略如图 2 所示。

在拥堵路段中，跟驰车辆的加速度 a 和转向角 δ 通常受前车运动状态所影响，因而需要基于前车状态来建立跟驰车辆的控制策略。在采集到前后车辆的速度差、相对距离及后车横向偏差、相对偏航角之后，利用深度确定性策略梯度算法，将车辆跟驰问题转换转换为特定奖励函数下的马尔可夫决策过程，并通过深度强化学习 agent 与车辆跟驰环境交互迭代学习，得到跟驰车辆的横纵向控制策略，也就是跟驰车辆的加速度和转向角，以此来调整跟驰车辆的运动状态，实现对跟驰车的最优控制[14]。

2.1　DDPG 算法原理

深度强化学习包含不断观察和奖励并与环境互动的智能体(agent)，以及因 agent 所采取的行动而发生变化的环境两部分。深度 Q 网络适用于具有少量离散输出的模型，但在连续动作空间中有可能决策失败。该研究利用在连续控制领域中表现良好的深度确定性策略梯度(DDPG)算法[15]，来学习 Actor 网络和 Critic 网络，将用于决策加速度和转向角的跟驰车作为 agent，其主要目标是最大化奖励函数。Actor 网络主要负责策略的生成，即根据跟驰车和前车的速度、相对速度以及相对距离来输出跟驰车的加速度，根据横向偏差和相对偏航角得到跟驰车方向盘的转角。Critic 网络则负责策略的改进，根据状态-动作对来输出 $Q(s_t,a_t)$，并按性能改进方向来更新 Actor 的策略参数。

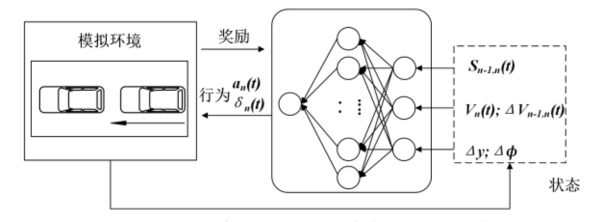


图 2　基于深度强化学习的车辆跟驰总体框图

Fig. 2　Deep reinforcement learning for car following

图 3 的 Actor 和 Critic 的网络架构由输入层，输出层和包含多个神经元的隐藏层所构成。DDPG 首先初始化缓冲区，Actor、Critic 网络参数 θ^μ、θ^Q 以及 Actor、Critic 相应的目标网络权重参数 $\theta^{\mu'}$、$\theta^{Q'}$，然后在每个训练周期中，根据 Actor 的策略 $a_t = \mu(s|\theta^\mu) + N_t$ 计算跟驰车的加速度和转向角，接下来观察奖励值 r_t 和下一时刻状态值 s_{t+1}，在得到奖励函数和状态值后利用 Critic 网络对当前状态 s_t 所采取的动作 a_t 进行评估，根据损失函数 L 来更新 Critic 的网络参数 θ^Q，并使用策略梯度来更新 Actor 的网络参数 θ^μ，最后根据 Actor 和 Critic 的网络权重的更新方向来更新目标网络权重 $\theta^{\mu'}$ 和 $\theta^{Q'}$，以此过程来不断优化 Actor 和 Critic，直到收敛。在这个过程中，DDPG 算法[15]的优化目标为

$$a = \underset{a=\mu(s|\theta^\mu)+N_t}{\arg\max}\ Q(s,a|\theta^Q) \tag{3}$$

Critic 网络依据损失函数(4)更新 θ^Q。

$$\begin{cases}y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta^{\mu'})|\theta^{Q'}) \\ L = \mathrm{E}(Q(s_t,a_t)|\theta^Q) - y_t)^2\end{cases} \tag{4}$$

Actor 网络使用策略梯度(5)来更新 θ^μ。

$$\nabla_{\theta^\mu} J \approx \nabla_a Q(s,a|\theta^Q)|_{s_t,\mu(s_t)} \nabla_{\theta^\mu}\mu(s_t|\theta^\mu)|_{s_t} \tag{5}$$

在 k 次优化后，采用策略(6)来更新 Actor 和 Critic 的目标网络参数 $\theta^{\mu'}$ 和 $\theta^{Q'}$。

$$\begin{cases}\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'} \\ \theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}\end{cases} \tag{6}$$

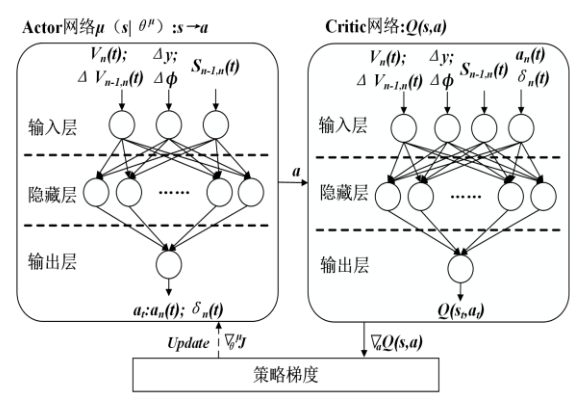


图 3　Actor 和 Critic 网络结构图

Fig. 3　Architecture of the actor and critic networks

2.2　车辆跟驰误差

跟驰车在跟随前车运动过程中，需要依据前车行驶状态和行驶轨迹，来规划跟驰车的控制策略。为了跟踪前车速度和行驶轨迹，表征跟驰车和前车之间的位置和速度及行驶轨迹之间的关系，使跟驰车保持一定的安全距离并沿期望路径行驶，建立横纵向联合控制的车辆跟驰误差模型如图 4 所示。

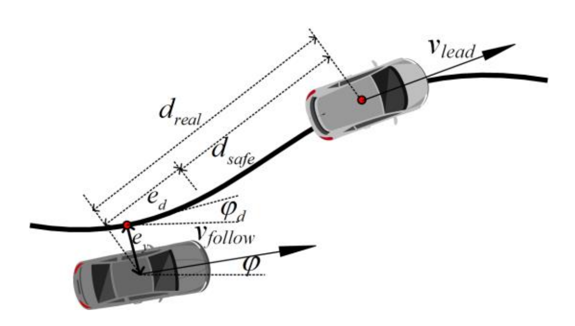


图 4 车辆跟驰误差模型

Fig. 4　Car following error model

在车辆纵向运动中，跟驰车需根据车辆速度和与前车的距离决策其加速度，从而安全、有效地跟随前车行驶。微观驾驶员行为的安全距离模型[16]为式(7)。

$$d_{safe}=\lambda_1(v_{follow}^2-v_{lead}^2)+t_{RT}*v+d_0 \tag{7}$$

其中：参数 λ_1 由车辆的最大制动能力决定；t_{RT} 是恒定车头时距(s)，一般为驾驶员反应时间；$\boldsymbol{v}$ 是跟驰车速度(m/s)；d_0 是跟驰车车速为零时的最小安全距离(m)。因车辆在跟驰行驶过程中，前车和后车速度差异小，将 $\lambda_1(v_{follow}^2-v_{lead}^2)$ 项忽略，故基于固定车间时距算法[16]，设计安全距离模型如(8)所示。

$$d_{safe}=t_{RT}*v+d_0 \tag{8}$$

根据两车的相对运动关系，结合车间相对速度和车间距误差定义车辆跟驰纵向模型(9)以直观反映跟随模式下跟驰车和前车的行驶状态。

$$\begin{cases}e_v=v_{follow}-v_{lead}\\ e_d=d_{real}-d_{safe}\end{cases} \tag{9}$$

上式中 $\boldsymbol{e_v}$ 为跟驰车速度 $\boldsymbol{v_{follow}}$ 和前车速度 $\boldsymbol{v_{lead}}$ 之间的差值，$\boldsymbol{e_d}$ 为实际距离 d_{real} 与安全跟车距离 d_{safe} 之间的差值。

在车辆横向运动中，跟驰车需要根据与行驶轨迹间的相对位置关系来获取横向速度和横摆角速度，并通过调整车辆转向角来保证有效地进行横向跟踪，减小横向偏差和相对偏航角[17]。

$$\begin{cases}\dot{e}_y=v_y+v_x e_\varphi\\ \dot{e}_\varphi=\dot{\varphi}-\dot{\varphi}_d=\dot{\varphi}-k_d v_x\end{cases} \tag{10}$$

其中：$\boldsymbol{e_y}$ 为横向位置偏差；e_φ 为偏航角误差；k_d 为跟驰轨迹曲率。

2.3　奖励函数设计

在强化学习中，奖励是环境对智能体动作的反馈和评估动作好坏的信号，通常为标量。在跟驰车横纵向联合控制中，需要用奖励函数来决定横向轨迹的跟随和纵向速度的控制。

跟驰控制问题可转换为以跟踪性、安全性和舒适性[5]等多个目标的优化问题。为了使跟驰车尽可能接近目标路径，同时保持更好的速度响应和稳定的加速行为，因此使用纵向速度误差 $\boldsymbol{e_v}$、距离误差 $\boldsymbol{e_d}$、横向偏差 $\boldsymbol{e_y}$、加速度 $\boldsymbol{a}$ 和转向角 δ 作为奖励函数的特征。另外还需考虑相对速度小于零、跟驰速度过小以及横向偏差过大所造成的异常情况给予模型惩罚 m。因此设计奖励函数(11)。

$$r_{total}=r_{follow}+r_{comfortable}+r_{safe}+m \tag{11}$$

为了有效地跟踪前车，基于纵向速度误差和横向偏差设计奖励函数(12)以实现纵向速度跟踪和横向路径跟踪。跟驰速度误差和横向偏差越小，r_{follow} 将越大，而当横向偏差小于 0.1m、跟驰速度误差小于 1m/s 时进行正向奖励 H，以更精准的进行路径跟踪与速度跟踪。

$$r_{follow}=-(\omega_1 e_v^2+\omega_2 e_y^2)+H \tag{12}$$

针对舒适性，基于加速度和转向角设计奖励函数(13)。跟驰加速度和转向角越小，横纵向的跟驰将会越稳定，舒适性也越好。

$$r_{comfortable}=-(\omega_3 a^2+\omega_4\delta^2) \tag{13}$$

针对安全性，基于安全距离误差设计奖励函数(14)。

$$r_{safe}=-\omega_5 e_d^2 \tag{14}$$

另外，设置终止条件为$|\boldsymbol{e_y}|>1$,$\boldsymbol{v}<0.5$,$d<0$，当触及终止条件时给模型惩罚 m，以防止横向偏差过大、跟驰速度过小，避免碰撞。

3　实验结果与分析

3.1　模型训练

基于 DDPG 的跟驰车横纵向联合控制的重点是模型特征选择和融合。由于使用驾驶员视角的视觉图像作为模型输入的可解释性太差，甚至导致神经网络无法学习到有用信息，因此本模型使用环境的特征向量 $\boldsymbol{X_{input}}$ 作为模型输入。

$$X_{input}=(\textstyle\int e_v,e_v,v,\int e_y,e_y,\dot{e}_y,\int e_\varphi,e_\varphi,\dot{e}_\varphi) \tag{15}$$

该模型根据当前的策略输出跟驰车的加速度 $\boldsymbol{a}$ 和转向角 δ 并更新其位置和速度。同时，环境更新前车的状态并返回当前步骤的奖励及更新策略。

网络训练步骤如下：

依据动作空间、状态空间及奖励函数，设计跟驰问题的深度强化学习要素；

初始化 Actor-Critic 网络，重置环境；

从环境中获取观测量 s_0，计算初始动作 $a_0=\mu(s_0)$，然后将初始动作设置为当前动作 $(a\leftarrow a_0)$,将当前的观测值设置为初始观测值 $(s\leftarrow s_0)$；

将动作 a 应用到环境中，获取下一时刻的观测值 s' 和奖励 r，然后从经验集中学习 (s,a,r,s')，计算下一时刻的动作 $a'=\mu(s')$，最后将 $(a\leftarrow a')$，来更新当前的动作，将 $(s\leftarrow s')$ 来更新当前的观测值;

以此循环，当达到终止条件，训练终止。

DDPG 算法的超参数如表 1 所示。

表 1　DDPG 算法超参数

Tab. 1　Super parameters of DDPG algorithm

参数	取值
Actor 网络学习率	0.0001
Critic 网络学习率	0.001
内存批大小尺寸	64
经验回放存储池	10^6
最大回合数	10^4
每一回合最多步数	600
时间步长	0.1

图 5 所示为训练跟驰模型中的奖励变化趋势，该训练从跟驰数据集中选取了 40 个跟驰事件的数据进行训练，其中红色曲线为各训练回合的平均奖励值，蓝色曲线为每个训练回合的奖励值，黄色曲线为每个回合开始时，critic 网络对折扣长期奖励的估计。平均奖励越高，跟驰的效果越好。实验经历 3548 个回合的训练，从图 5 中可明显的看出在约 1400 个回合时，奖励函数逐步收敛。

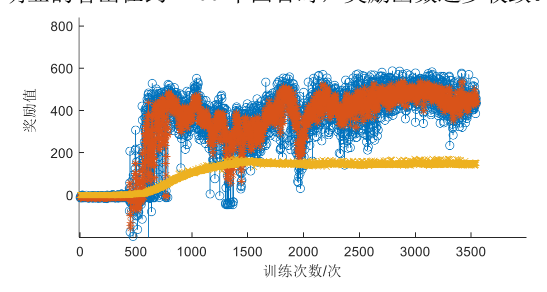


图 5　跟驰模型训练奖励值整体变化图

Fig. 5　Car following model training reward value

在图 5 中，最后 100 个回合的奖励值如图 6 所示，可以看出该算法稳定有效。

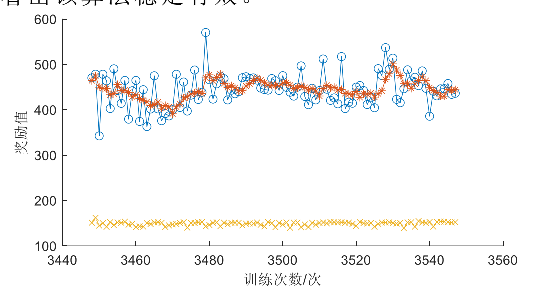


图 6　跟驰模型训练最后 100 个回合奖励值

Fig. 6　The last 100 rounds of the car following model training

3.2　模型测试

利用 MATLAB/Simulink 搭建跟驰控制仿真系统，建立整车动力学模型，整车动力学参数如表 2 所示。

表 2　整车动力学参数

Tab. 2　Vehicle dynamics parameters

参数名	参数	参数值
整车质量	m	1600kg
车身绕 z 轴转动惯量	I_z	2875kg/m^2
质心到前轴距离	l_f	1.4m
质心到后轴距离	l_r	1.6m
前轮胎转弯刚度	C_f	19000N/rad
后轮胎转弯刚度	C_r	33000N/rad

基于著名的 NGSIM 真实驾驶员跟驰驾驶数据集对跟驰控制策略进行测试与验证，从 I-80 路段的车辆轨迹数据中提取 1341 个跟驰事件，每个跟驰事件包含前车速度、跟驰车速度、两车相对速度以及两车相对间距，持续时间在 15s 以上。如表 3 所示为某一个跟驰事件的部分数据。

表 3　跟驰数据结构

Tab. 3　Car following data structure

跟车间距(m)	跟驰车速(m/s)	相对速度(m/s)	前车速度(m/s)
9.995	7.773	2.368	10.141
10.197	8.015	2.015	10.030
10.367	8.231	1.701	9.932
……	……	……	……

调整加速度和转向角使跟驰车跟随前车沿着曲率变化道路行驶，在满足安全距离、速度、加速度和转向角的限制条件下，计算最优跟驰动作。另外根据车辆动力学物理限制，设置跟驰车加速度 $\boldsymbol{a}$ 和转向角 δ 范围[5]如(16)所示。

$$\begin{cases} -3m/s^2 \le a \le 2m/s^2 \\ -0.5rad \le \delta \le 0.5rad \end{cases} \tag{16}$$

从 1341 对跟驰对数据中，随机选择一组跟驰数据验证本文所提跟驰决策方案并与 MPC 跟驰方案进行对比。实验中的车辆间距、速度及加速度变化如图 7~9 所示。设置前车和跟驰车的初始车间距为 15m，跟驰车纵向速度为 6.7m/s，前车纵向速度为 6.1m/s。图 7 为跟车间距图，可以看出，跟驰车和前车的车距始终保持相对稳定，DDPG 算法的跟车距离整体上比人类驾驶员和 MPC 决策的距离小，实现了高效跟驰。

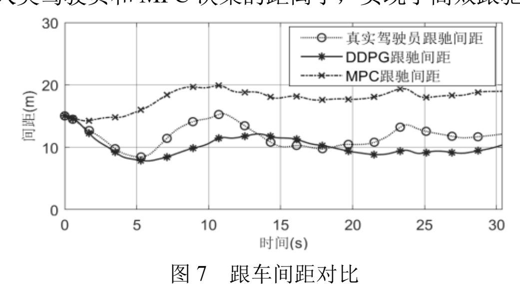


图 7　跟车间距对比

Fig. 7　Comparison of the car following distance

图 8 为前车速度，真实驾驶员跟驰速度、DDPG 算法所决策的跟驰速度、MPC 算法所决策的跟驰速度。从图中可以看出前车经历了先加速再减速然后匀速行驶的过程，而 DDPG 算法在跟随行驶中的速度变化更加稳定。

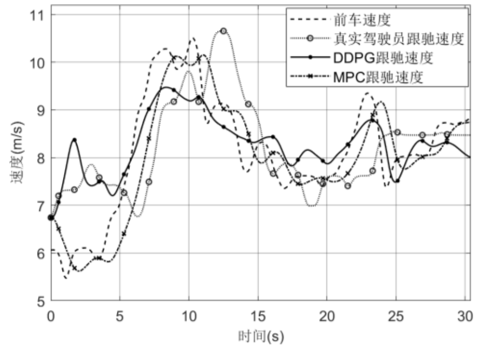


图 8　车辆速度变化

Fig. 8　Vehicle speed

图 9 为跟驰车加速度曲线图，可以看出，跟驰车通过合理调整加速度有效地调整了车速与间距，且保持了较为平稳的加速度。

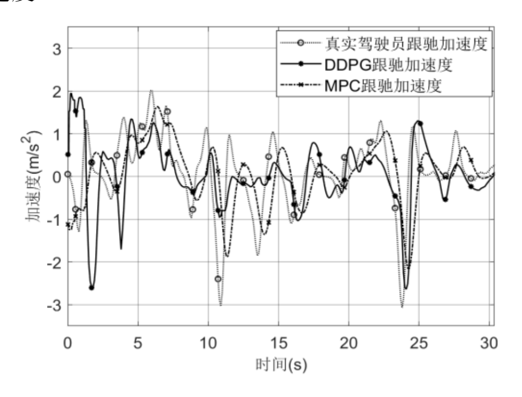


图 9　跟驰车加速度变化

Fig. 9　Car following acceleration

为评价本文所用算法的跟驰效果，采用平均绝对误差(17)作为模型评价指标。

$$MAD=(\sum_{i=1}^{n}|y_i-\bar{y}_i|)/n \tag{17}$$

其中：y_i 为单个观测值；$\bar{y}_i$ 为算数平均值。表 4 给出了 DDPG 算法、MPC 算法和真实驾驶员实测值的平均绝对误差。

表 4　模型误差

Tab. 4　Model error

MAD	真实驾驶员	MPC	DDPG
加速度	0.5722	0.5507	0.5407
速度	0.8089	0.9823	0.4932
距离	1.5306	1.3390	1.4408

从表 4 可以看出，DDPG 算法的加速度平均绝对误差最小，实现了更加平稳舒适的跟驰效果；DDPG 算法的速度误差相对于人类驾驶员和 MPC 算法都有所降低，算法跟驰有效且具有较强的自适应能力，能够保持更加稳定的行驶速度；DDPG 算法决策的跟车间距平均绝对误差比真实驾驶员小，故 DDPG 算法的决策效果更加稳定且能保持较小的车距。

图 10 为前车行驶的道路轨迹曲率，图 11 为横向控制实验结果。其中，跟驰车的初始横向偏差设置为 0.2m，初始偏航角设置为-0.1rad。跟驰车通过控制转向角，使得偏航角误差快速减小，由于道路轨迹曲率不断变化，故跟驰车不断微调转向角，使横向偏差和相对偏航角较小，DDPG 算法和 MPC 算法的横向控制效果基本一致。

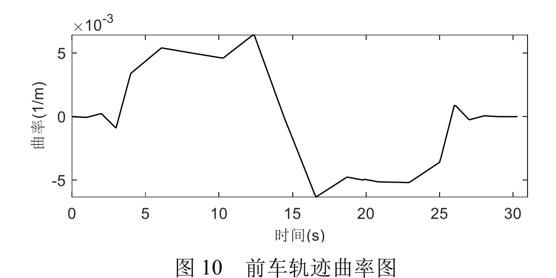


图 10　前车轨迹曲率图

Fig. 10　Front car trajectory curvature

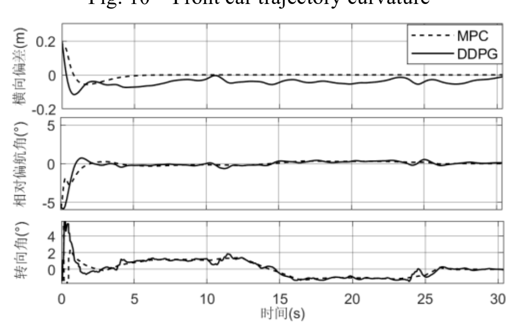


图 11　横向控制效果图

Fig. 11　Lateral control results

为了评价跟驰模型的预测性能，选取加速度变化率(jerk)来评价跟驰舒适性，选取车头时距(thw，time headway)评价跟驰安全性和有效性。在跟驰行驶中的车头时距通常保持在1s~4s 内，车头时距越小跟踪越紧，跟驰效率越高，但若低于1s 容易发生碰撞，而大于 4s 则通常不属于跟驰行驶。

$$\begin{cases} jerk=\sum_{t=0}^{n}\left|\Delta a/\Delta t\right| \\ thw=\left|\frac{1}{n}\sum_{t=0}^{n}d\Big/\frac{1}{n}\sum_{t=0}^{n}v\right| \end{cases} \tag{18}$$

从表 5 可以看出，DDPG 算法所决策跟驰车的加速度变化率比真实驾驶员决策的加速度变化率小，可保证车辆跟驰的舒适性，避免频繁加减速带来的不适感。另外，跟驰车的车头时距保持在了 1-4s 的安全车头时距范围内，比真实驾驶员和 MPC 算法决策的车头时距更小，跟驰效率更高。

表 5　跟驰模型评价

Tab. 5　Car following model evaluation

	真实驾驶员决策	MPC 决策	DDPG 决策
jerk(m/s^3)	531.4185	299.3043	426.4647
thw(s)	1.4610	2.2279	1.2624

为验证多种跟驰工况下的实验效果，另在 NGSIM 数据集中随机选择一个跟驰事件，实验结果如图 12~15 和表6~7 所示。可以看出，在该跟驰事件下，DDPG 算法决策的加速度及其变化率更小，跟驰车速度更加稳定，车头时距保持在 1s~4s 的范围内，依然实现了更加舒适、稳定和安全地跟驰。

4　结束语

本文基于车辆三自由度动力学模型并结合横纵向联合控制，搭建了一种车辆跟驰控制模型，然后基于深度强化学习DDPG 算法建立了决策模型以对跟驰车的加速度和转向角进行决策，以确保跟驰车辆的行驶安全有效和舒适。实验利用人类驾驶数据集 NGSIM 对模型进行了训练测试与评估，并与 MPC 跟驰控制算法相比，结果表明本文所用方法在保证安全的情况下，跟驰距离和加速度变化率更小，比人类驾驶员表现更佳，对保证交通安全和提升道路通行能力具有较大意义。

目前，跟驰控制功能相对独立，如将车辆跟驰控制与车道保持辅助系统和车道变换辅助系统相结合，将实现更高层次的自动驾驶控制。

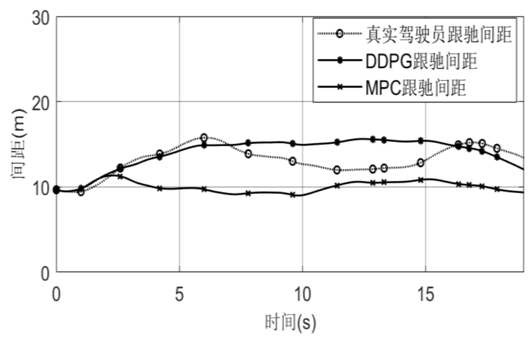


图 12　跟车间距对比

Fig. 12　Comparison of the car following distance

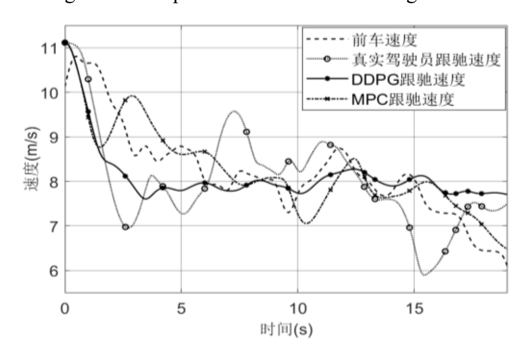


图 13　车辆速度变化

Fig. 13　Vehicle speed

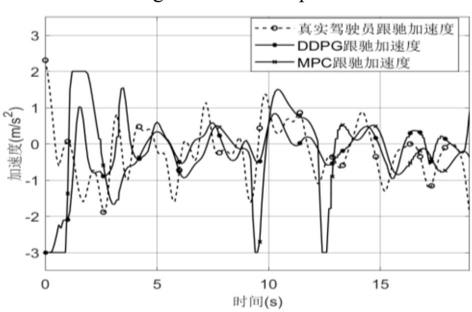


图 14　跟驰车加速度变化

Fig. 14　Car following acceleration

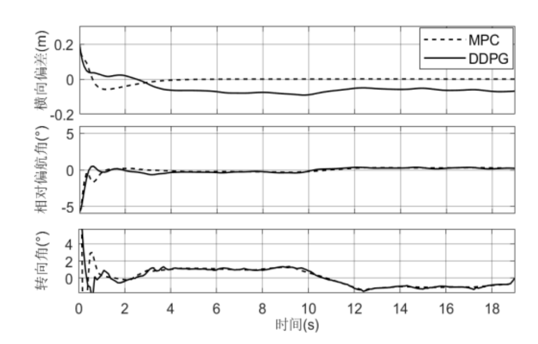


图 15　横向控制效果图

Fig. 15　Lateral control results

表 6　模型误差

Tab. 6　Model error

MAD	真实驾驶员	MPC	DDPG
加速度	0.5766	0.8241	0.5161
速度	0.8630	0.7859	0.4370
距离	1.4368	0.4773	1.3314

表 7　跟驰模型评价

Tab. 7　Car following model evaluation

	真实驾驶员决策	MPC 决策	DDPG 决策
jerk(m/s^3)	408.2924	337.4202	265.5867
thw(s)	1.6712	1.2306	1.7247

参考文献：

[1] Saifuzzaman M, Zheng Z. Incorporating human-factors in car-following models: a review of recent developments and research needs [J]. Transportation research part C: emerging technologies, 2014, 48: 379-403.

[2] Pipes L A. An operational analysis of traffic dynamics [J]. Journal of applied physics, 1953, 24 (3): 274-281.

[3] 张兰芳，朱佩玄，杨旻皓，等. 基于数据驱动的城市地下快速路跟驰行为模型构建 [J]. 同济大学学报：自然科学版, 2021, 49 (05): 661-669. (Zhang Lanfang, Zhu Peixuan, Yang Minhao, *et al.* Modeling of car-following behavior on urban underground expressways based on data-driven Methods [J]. Journal of Tongji University: Natural Science, 2021, 49 (05): 661-669.)

[4] 罗颖，秦文虎. 基于 IDM 与 RBFNN 的组合型车辆低速跟驰模型 [J]. 计算机应用研究, 2019, 37 (8): 1-7. (Luo Ying, Qin Wenhu. Combination low-speed car-following model based on IDM and RBFNN [J]. Application Research of Computers, 2019, 37 (8): 1-7.)

[5] Zhu Meixin, Wang Yinhai, Pu Ziyuan, *et al.* Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving [J]. Transportation Research Part C: Emerging Technologies, 2020, 117: 102662.

[6] Pan Feng, Bao Hong. Preceding vehicle following algorithm with human driving characteristics [J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2021, 235 (7): 1825-1834.

[7] 朱冰，蒋渊德，赵健，等. 基于深度强化学习的车辆跟驰控制 [J]. 中国公路学报, 2019, 32 (06): 53-60. (Zhu Bing, Jiang YuanDe, Zhao Jian, *et al.* A car-following control algorithm based on deep reinforcement learning [J]. China Journal of Highway and Transport, 2019, 32 (06): 53-60.)

[8] Camacho E F, Alba C B. Model predictive control [M]. Springer science & business media, 2013.

[9] Hu Xiaosong, Zhang Xiaoqian, Tang Xiaolin, *et al.* Model predictive control of hybrid electric vehicles for fuel economy, emission reductions, and inter-vehicle safety in car-following scenarios [J]. Energy, 2020, 196: 117101.

[10] Mao Jin, Yang Lei, Hu Yuanbo, *et al.* Research on Vehicle Adaptive Cruise Control Method Based on Fuzzy Model Predictive Control [J]. Machines, 2021, 9 (8): 160.

[11] Muhrer E, Vollrath M. The effect of visual and cognitive distraction on driver’s anticipation in a simulated car following scenario [J]. Transportation research part F: traffic psychology and behaviour, 2011, 14 (6): 555-566.

[12] 许芳，张君明，胡云峰，等. 智能车辆路径跟踪横纵向耦合实时预测控制器 [J]. 吉林大学学报：工学版, 2021, 51 (6): 2287-2294. (Xu Fang, Zhang Junming, Hu Yunfeng, *et al.* Lateral and longitudinal coupling real-time predictive controller for intelligent vehicle path tracking [J]. Journal of Jilin University: Engineering and Technology Edition, 2021, 51 (6): 2287-2294.)

[13] Wang Hong, Huang Yanjun, Khajepour A, *et al.* Crash mitigation in motion planning for autonomous vehicles [J]. IEEE transactions on intelligent transportation systems, 2019, 20 (9): 3313-3323.

[14] Zhu Meixin, Wang Xuesong, Wang Yinhai. Human-like autonomous car-following model with deep reinforcement learning [J]. Transportation research part C: emerging technologies, 2018, 97: 348-368.

[15] Lillicrap T P, Hunt J J, Pritzel A, *et al.* Continuous control with deep reinforcement learning [J]. arXiv preprint arXiv: 1509. 02971, 2015.

[16] Puan O C, Mohamed A, Idham M K, *et al.* Drivers behaviour on expressways: headway and speed relationships [C]// IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2019, 527 (1): 012071.

[17] Wang Yulei, Ding Haitao, Yuan Jinxin, *et al.* Output-feedback triple-step coordinated control for path following of autonomous ground vehicles [J]. Mechanical Systems and Signal Processing, 2019, 116: 146-159.